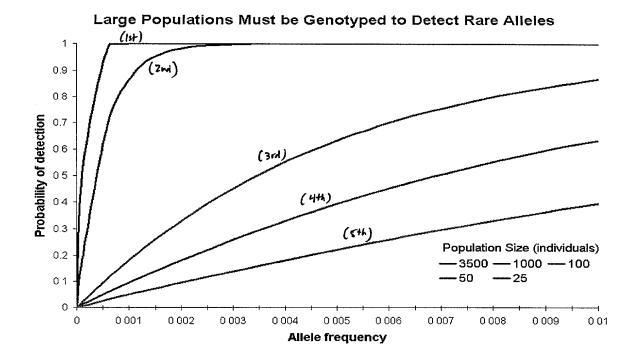
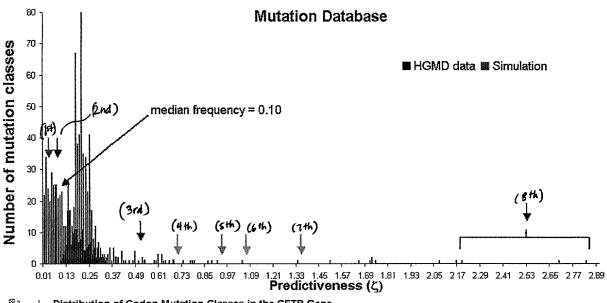
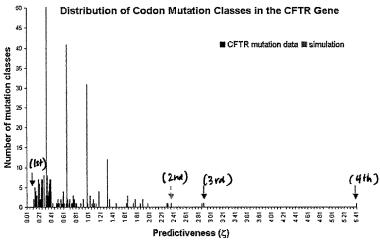
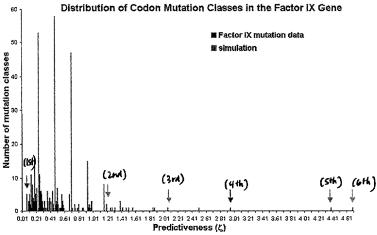
FIGURE 1

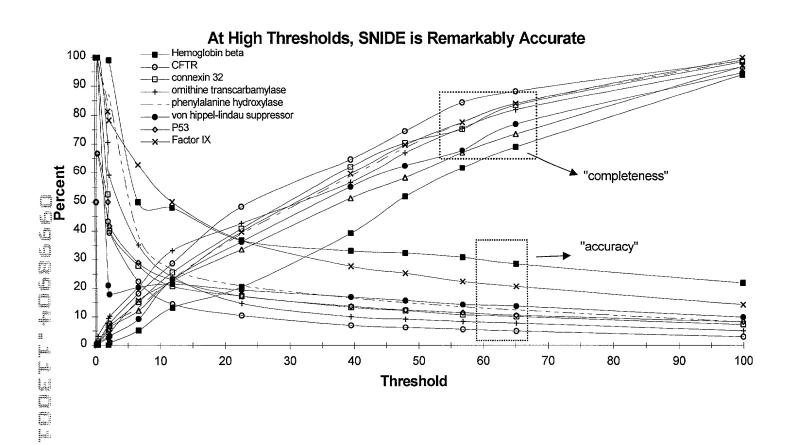


## FIGURE 2

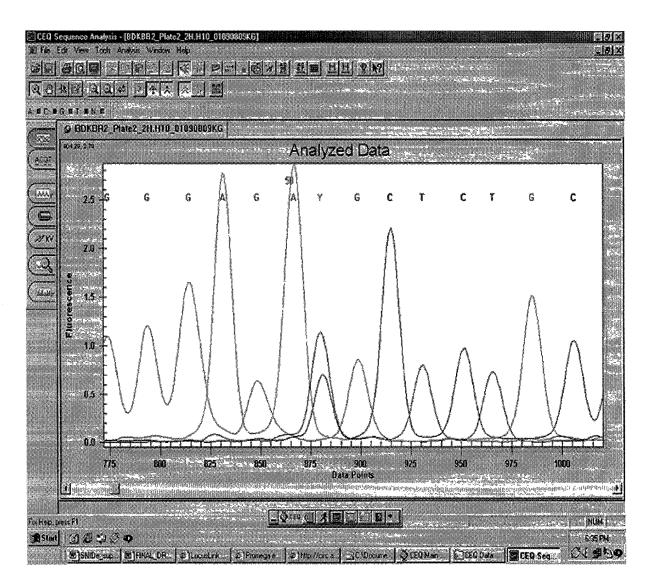








## FIGURE 4



## FIGURE 5

## Matrix construction

Gather variation data collection with property of interest (e.g., diseasecausing, high frequency, etc.) Determine predictiveness scores from relative frequencies of variations Modify predictiveness scores based on other considerations (e.g., codon usage, structure, etc) Evaluate query sequence(s) against predictiveness matrix to identify likely

Matrix deployment

variants have desired properties (e.g., disease-causing, frequency, etc)